

UM ESTUDO COMPARATIVO DA DESAMBIGUAÇÃO LEXICAL REALIZADA POR TRADUTOR ELETRÔNICO E TRADUTORES HUMANOS

Adriana Riess Karnal¹

Resumo

Este estudo trata da desambiguação lexical realizada por tradutores eletrônicos e humanos. Sabe-se que há muitas ambiguidades (polissemia e homonímia) nas línguas e a escolha por um significado ou outro passa a ser um problema para os tradutores. Assim, é objetivo do artigo analisar como os tradutores tanto humanos quanto as máquinas realizam a desambiguação a partir dos modos distintos de raciocínios envolvidos entre os dois. Por um lado, a pesquisa está fundamentada na discussão de Bevilacqua e Reulliard (2013) que explica as competências necessárias da tradução humana, por outro lado, a investigação sobre a tradução pela máquina baseia-se nas metodologias que essa área utiliza descritas por Mitkov (2003) e Hutchins (2003). Nesse sentido, comparam-se duas traduções, uma realizada pelo Google e outra de A. Brown do poema *A Pera*. Analisou-se o poema a partir das escolhas lexicais feitas quando a desambiguação foi necessária. Os resultados mostram que a tradução pela máquina está próxima da tradução humana, isso leva a crer que há um avanço evidente nesse tipo de tradutor mesmo para textos metafóricos, como é a poesia.

Palavras-chave: Tradução. *Machine translation*. Tradução eletrônica. Desambiguação lexical. Polissemia.

Abstract

This study deals with lexical disambiguation carried out by electronic and human translators. It is known that there are many ambiguities (polysemy and homonymy) in languages and choosing the most suitable meaning may be a problem for translators. Thus, it is the objective of the article to analyze how both human translators and machines disambiguate according to the different reasoning involved between the two. On the one hand, the research is based on Bevilacqua and Reulliard (2013) that explain the necessary competences of human translation, on the other hand, Mitkov (2003) and Hutchins (2003) base investigation on machine translation on the methodologies that this area uses described. Bearing this in mind, two translations are compared; one performed by Google and the other by A. Brown from the poem *A Pera*. The poem was analyzed from the lexical choices made when disambiguation was necessary. Results show that machine translation is close to

¹ Doutora em Linguística Aplicada (PUCRS, Brasil) e bolsista DOCFIX do Programa de Pós-Graduação em Letras na Universidade de Santa Cruz do Sul (Brasil), Edital 04/2018 CAPES/FAPERGS. adrianariess@unisc.br

human translation, this makes us believe that there is an advance in this type of translator even for metaphorical texts, such as poetry.

Keywords: Translation; Electronic translation. Lexical disambiguation. Machine translation. Polysemy.

1. Introdução

É característica das línguas faladas no mundo que as palavras sejam polissêmicas ou homônimas, e a desambiguação lexical, de um modo geral, é decidida no contexto. A polissemia ou a homonímia, nesse sentido, passa a ser uma dificuldade para tradutores, porque optar por uma tradução de equivalência entre uma língua e outra depende de conhecimentos que não estão necessariamente no (co)texto. Se comparados os tipos de tradução entre aquele realizado pelas máquinas eletrônicas e o feito por tradutores humanos, as dificuldades são ainda maiores, porque os sistemas de tradução automática não são ainda capazes de reconhecer sutilezas linguísticas que contemplem as tantas situações das palavras em uso. Conforme Mitkov (2003), algumas questões de linguagem para tradutores eletrônicos que acarretam problemas são a desambiguação do sentido da palavra, a anáfora e o processamento sintático de modo geral. Este estudo investiga a área de *Machine Translation* com vista à tradução realizada pelo *Google Translate* e como essa ferramenta é capaz de desambiguar o léxico polissêmico. Ao mesmo tempo, tem-se o olhar voltado para traduções realizadas por humanos em casos de desambiguação lexical onde o autor precisa decidir qual a palavra mais adequada para a tradução.

Para explicar como a tradução feita pela máquina e a tradução de humanos diferenciam-se é preciso analisar os tipos de raciocínio em jogo. A tradução feita pela máquina funciona sob a perspectiva da inteligência artificial, e, muitas vezes, tenta simular o raciocínio humano, tal como as inferências. No que se refere ao tradutor do Google, Somers (2003) argumenta que o modelo de tradução estatística, no qual a ferramenta está fundamentada, não lança mão de dados linguísticos tradicionais; tudo o que faz é alinhar o texto-fonte em paralelo ao texto-alvo calculando as probabilidades de correspondências de uma palavra ou palavras e suas concomitantes na outra língua a partir de um *corpus* previamente traduzido. Por outro lado, o tradutor francês Henri Meschonnic (apud BASTOS, 2012) diz que a visão normalmente adotada pelos especialistas e profissionais da tradução é o da “estilística comparada” em que há noções de equivalência, fidelidade e transparência entre a língua de saída e língua de chegada, assim a tradução é interpretação.

A fim de analisar a desambiguação lexical faz-se uma comparação do poema *A Pera* de Vinícius de Moraes entre o tradutor do Google e uma tradução humana. O poema é rico em metáforas, mas em termos de ambiguidade, ela acontece pela fonologia das palavras *cera* e *ser a*, a primeira inicia o poema e a outra encerra. O que se observa é que as traduções são similares, embora na última frase há uma implicatura pragmática relacionada ao verbo mandar que não é tão transparente na tradução humana, e acaba por diferenciar as duas traduções.

O artigo está dividido em quatro partes. Na primeira seção discute-se a desambiguação lexical que é um dos problemas de linguagem a serem enfrentados na tradução, assim, uma pergunta de pesquisa é como humanos e máquinas

desambiguar. A segunda seção trata da tradução humana e descreve as competências necessárias para que o tradutor seja capaz de empreendê-las para uma tradução mais fiel da obra original. Na terceira seção aborda-se as metodologias que a área de *machine translation* utiliza na tradução, e como esse campo de estudos tem avançado, em especial a tradução estatística do Google. A quarta seção tem cunho prático, trata-se de uma comparação entre os tipos de traduções discutidas aqui através da análise do poema *A Pera*, já mencionado anteriormente. Finalmente, o trabalho se encerra com uma discussão sobre a desambiguação e como ela é resolvida pelas metodologias de tradução tanto humanas quanto às da máquina.

2. Desambiguação lexical

Ao trazer o assunto da desambiguação lexical, é preciso tomar consciência de que as palavras ambíguas em uma língua podem ser categorizadas em polissêmicas e homônimas. A polissemia tem como característica a existência de múltiplos significados para uma única palavra; já as palavras homônimas são aquelas que também carregam vários significados, entretanto, têm origem diferente, tendo convergido foneticamente. Um exemplo de polissemia é a palavra *tubo*, como em *o médico retirou o tubo* que tanto pode remeter ao significado de tirar o paciente do estado de coma, como a retirada de um tubo qualquer. Já na homonímia, o campo semântico das palavras é muito distante, tal como em *manga* referindo-se à fruta ou à parte da camisa. Uma das explicações para a existência de múltiplos significados para um mesmo significante se refere ao trabalho cognitivo de alimentação e manutenção do léxico mental. Sabe-se que os falantes tendem à “lei do menor esforço” ou ao princípio da economia, e isso implica um vocabulário capaz de transmitir o máximo de sentidos. Conforme Dela Rosa e Eskenazi (2011), o vocabulário de uma língua é uma mistura de palavras ambíguas e não ambíguas.

É necessário iniciar este estudo com a explicação sobre a polissemia e homonímia, pois estes são os fenômenos linguísticos que levam à desambiguação, nesse sentido, desambiguar é uma espécie de raciocínio que o leitor é capaz de realizar durante uma tarefa cognitiva como a tradução, por exemplo. Stevenson e Wilks (2003) explicam que a realização da desambiguação exige que o leitor reconheça o contexto no qual o significado da palavra se encontra. Para eles, a desambiguação é o processo de decidir os sentidos das palavras no contexto. A ambiguidade linguística é também tratada pela área de estudos em leitura, muitos autores explicam o fenômeno pelas sentenças *Garden path* (JUFFS e HARRINGTON, 1996; PERFETTI, 2001; STANOVICH, 1980). Essas sentenças produzem um atraso no processamento porque o leitor deve escolher qual caminho de compreensão irá tomar, e isso traz consequências para a memória de trabalho necessária para a manutenção da informação lida, além de fazer uma espécie de rompimento ao integrar a palavra ao contexto que faz fluir a leitura, também necessária para o acesso do significado.

Um dos problemas de tradução encontrados nas sentenças *garden path* pode ser visto no exemplo em inglês *I convinced her children are noisy* (*eu a convenci que as crianças são barulhentas*), esta frase pode ser compreendida mais facilmente somente quando o leitor perceber que o pronome relativo *that* vem depois de *her*. Também pode ocorrer a compreensão pela entonação mais forte em *children*. Os estudos sobre esse tipo de sentença resultam em contextos desviantes que

demonstram como nossa mente/cérebro não passa diretamente das palavras para o significado. Nesse caso, o processamento sintático é indispensável para a configuração do significado, assim, cabe ao tradutor ter domínio sintático suficiente da língua, de tal modo que saiba desambiguar não apenas palavras isoladas, mas uma sentença *Garden path*, por exemplo.

Há vários modelos cognitivos que tratam da desambiguação, tais como os conexionistas descritos por Waltz e Pollack (1985), até as abordagens baseadas puramente em dicionário, como o de Lesk (1986). Aqui, acredita-se que um modelo lexical deve explicar a desambiguação baseado tanto nas habilidades cognitivas mais básicas, tal como o reconhecimento das palavras (por exemplo, o tradutor deve saber que *disc* e *disk* são a mesma palavra) até as habilidades mais altas, em que a capacidade inferencial é fundamental. Assim, os modelos situacionais (que contemplam o contexto) parecem os mais adequados para explicar a desambiguação. De fato, uma metodologia de tradução eletrônica mais recente fundamentada no conceito de ontologia tem demonstrado ser eficiente.

3. Tradução humana

É, em certa medida, ingênuo pensar que para fazer a tradução de uma língua para outra basta ser um falante bilíngue, a tradução exige conhecimentos tanto linguísticos, culturais quanto estéticos. Conforme afirmam Bevilacqua e Reulliard (2013) há uma competência tradutória que precisa ser desenvolvida, para elas, transformar um bilíngue em tradutor é desenvolver competências extralinguística, instrumental, de conhecimentos de tradução e também uma competência estruturante, chamada de estratégica. O conhecimento bilíngue, é claro, é necessário, porque está estruturado nas línguas materna e estrangeira, em todos seus aspectos gramaticais, textuais, fraseológicos, pragmáticos (diferentes usos) e sociolinguísticos (a norma e suas variantes). Ao mesmo tempo, é preciso saber que o processo tradutório envolve a unidade de tradução, equivalência e fidelidade, adaptação, função, métodos e estratégias.

Diferentemente da máquina, o tradutor humano deve ter consciência linguística, é isso que o capacita a resolver problemas quando se depara com o que tradicionalmente se chama de desvios de linguagem. No caso deste trabalho, a desambiguação lexical é um desvio da linguagem, e é uma tarefa cognitiva fundamental na tradução, porque o tradutor deve escolher entre uma palavra ou outra diante da ambiguidade. A escolha de significados passa a ser um dos grandes problemas porque frequentemente não há equivalências entre uma língua e outra. A pesquisa sobre leitura e ciência cognitiva tem mostrado que o acesso ao léxico mental exige tanto da memória de longa duração (a semântica, por exemplo) quanto da memória de trabalho, aquela utilizada durante a tradução. Assim, quando a desambiguação é necessária, será exigido do tradutor uma espécie de memória auxiliar, já que nem sempre é possível acessar por exemplo, palavras raras ou idiossincrasias que são pouco frequentes e não estão disponíveis *on-line* (PERFETTI, 2007; DEHEANE, 2012).

A memória auxiliar da qual se fala aqui é, de fato, parte da competência instrumental da tradução, pois atualmente o tradutor conta com muitos recursos tecnológicos que o ajudam, por exemplo, os softwares de tradução, dicionários bilíngues, dicionários especialistas, bem como corpora eletrônicos. Os dicionários

bilíngues mais tradicionais, tal como o Michaelis, estão disponíveis na web, entretanto, nem sempre resolvem uma ambiguidade, porque ficam no nível descritivo da palavra e não oferecem contexto. Entretanto, dicionários especializados baseados em ontologia, acabam por refinar o significado, porque funcionam no nível pragmático. Esse é o caso do *Field*, dicionário sobre os termos do futebol lançado para a copa do mundo no Brasil.

As ferramentas descritas no parágrafo anterior, tais como dicionários e corpora eletrônico, dizem respeito à importância da Tradução Automática (TA) para o trabalho dos tradutores humanos (profissionais) e a necessidade da pós-edição. Essas ferramentas encontram-se embutidas em vários sistemas de Memórias de Tradução. Vale ainda ressaltar, que os sistemas de tradução não foram primeiramente pensados para o texto poético, e se adaptam muito mais aos textos técnico-científicos porque são menos metafóricos.

Em relação à competência estética, dentre as discussões que se tem na área, a pergunta que permanece é quanto à função do tradutor, ele é um intérprete ou criador de linguagem? No caso do tradutor como intérprete seu trabalho é manter a fidelidade ao texto original da língua de partida. Por outro lado, o tradutor pode ser visto como um criador de linguagem, principalmente em certos gêneros textuais, tal como o poético, que utiliza muitos desvios de linguagem, como a própria ambiguidade lexical, a fim de provocar sentidos no leitor. Vale lembrar que em muitas situações não há equivalência do léxico de um a um, por exemplo, no poema *A Pera*, que será analisado mais adiante, não há uma tradução exata em inglês para o termo *será* como expressão de dúvida e não o verbo *ser* no futuro. Talvez a tradução para *será?* possa ser a expressão coloquial *really*, mas mesmo ela é ambígua e dependente do contexto. Nessa situação, o autor precisa lançar mão de seu próprio léxico mental e de sua capacidade criativa para traduzi-la. No texto ficcional ocorre o mesmo, nos livros da personagem Harry Potter o nome do jogo *Quadribol*, que originalmente é *Quidditch*, exige que o tradutor invente um nome, já que tal palavra inexistente em qualquer língua.

4. Tradução pela máquina

Machine Translation ou Tradução pela Máquina é uma área da pesquisa em processamento da linguagem natural (PLN), sendo que a pesquisa sobre tradutores eletrônicos está inserida aqui. Esse campo do conhecimento tenta responder à computação como a máquina pode compreender uma linguagem natural para, então, realizar a tradução. Como linguistas, sabemos dos inúmeros problemas de linguagem enfrentados pelos tradutores, a tradução pela máquina é mais problemática ainda, pois, obviamente, não pensa como humanos. Algumas questões de linguagem (MITKOV, 2003) que acarretam problemas são a desambiguação do sentido da palavra, a anáfora e o processamento sintático de modo geral. Além dos problemas linguísticos descritos acima, a máquina também apresenta dificuldades em traduzir metáforas, já que elas representam a expressão de culturas específicas e a própria criatividade da linguagem.

Um exemplo de metáfora na língua portuguesa é a expressão *ele é um bom garfo* que equivale a *ele é um comilão*, a compreensão disso exigiria um conhecimento cultural da relação entre *bom garfo* e *comer demais* que a máquina não tem. É por essa razão que a tradução acaba por ser literal perdendo seu sentido metafórico. Não seria novidade dizer que há problemas na tradução dos tradutores eletrônicos, tais

como o *Google translate*, mas esse é um esforço que a área tenta enfrentar. No caso desse trabalho, as palavras ambíguas (polissêmicas ou homônimas) são também difíceis de serem traduzidas, pois a máquina deve optar por um significado ou outro, tal como “*she waited for him by the bank*”, em que a palavra *bank* pode ser tanto *banco* quanto *margem de rio*.

Se, por um lado, a tradução feita por humanos lida com questões que vão desde o nível sintático (por exemplo, a definição de um tempo verbal mais apropriado, a preposição correta...) até o nível semântico-pragmático (decisão sobre o contexto adequado de uma palavra polissêmica, ou a lógica da ambiguidade lexical), a tradução pela máquina lança mão de metodologias específicas para a resolução dessas questões. Ao retomar o histórico, Hutchins (2003) descreve duas fases metodológicas que a área já vivenciou. A primeira é chamada de *Rule-Based Systems*, a segunda, e mais recente, é a Abordagem baseada em *Corpus* que faz uma contagem estatística da frequência da palavra traduzida.

A *Rule-Based Systems* compreendeu de três momentos, em sua primeira versão foi definida como *direct translation*. Nesse período, havia pouca análise sintático-semântica, muito similar a uma tradução rudimentar na qual a equivalência lexical era de um para um, trabalhando no nível do dicionário, com eventuais resoluções de ambiguidades, ordem de palavras ou expressões idiomáticas da língua-alvo. O segundo momento dessa abordagem é definido como interlíngua. Aqui, as representações sintáticas e semânticas comuns em uma ou mais línguas, conforme diz Hutchins (2003), estão em jogo. Isso implica dizer que há uma primeira tradução da língua-fonte para uma língua-alvo que não se refere a uma língua em particular, daí o título interlíngua. É a partir dessa interlíngua que será gerada a tradução da língua-alvo final (que pode ser mais de uma). A terceira abordagem da *Rule-Based Systems* é a *transfer* (transferência), resumidamente, ela transfere ou transforma as estruturas da língua fonte para a língua-alvo, como se utilizasse, pelo menos, três dicionários, um da língua-fonte, um da língua-alvo e outro bilíngue que relaciona os dois anteriores (HUTCHINS, 2003).

A abordagem descrita acima baseada nas regras da língua prevaleceu até os anos 80, conforme diz Hutchins (2003), mas não permitiu que traduções de excelência se concretizassem. A tradução direta muito pouco analisa todos os subsistemas linguísticos que a tradução exige; já a interlíngua não reflete cada diferença semântica inerente às línguas ali envolvidas. Quanto à transferência, ela não pressupõe o conhecimento do mundo real. Isso porque esta abordagem não prevê que o léxico das línguas são eles mesmos representantes de conhecimentos extralinguísticos. O autor ilustra essa afirmação com o exemplo do Francês *fleuve* e *rivière*, em que o primeiro se refere ao rio em si, e o segundo ao rio que deságua no mar. Não obstante, a abordagem não cessou, e atualmente se desdobra em sistemas mais híbridos, como a resolução de anáforas e o sistema baseado em conhecimentos.

Na década de 90, a abordagem mais utilizada em *machine translation* (doravante MT) foi a baseada em regras. Uma vez problematizadas as dificuldades iniciais da área da tradução pela máquina, uma nova abordagem aparece no início dos anos 2000, e parece ser mais bem-sucedida, a chamada *Tradução Estatística*. De fato, essa é a metodologia usada pelo tradutor eletrônico do Google, conforme descreve Somers (2003), essa metodologia depende de um *corpus* bilíngue, mas o procedimento depende da modelagem estatística da ordem das palavras e de equivalências de palavras da língua fonte para a língua-alvo. Nas palavras de Somers

(2003, p. 513): “there is thus a focus on the mathematical aspects of estimation of statistical parameters for the language models.” Com um olhar mais aproximado, a tradução estatística pode ser considerada como um novo paradigma da área. Para explicar isso, Somers (2003) argumenta que ela não lança mão de dados linguísticos tradicionais; tudo o que faz é alinhar o texto-fonte em paralelo ao texto-alvo calculando as probabilidades de correspondências de uma palavra ou palavras e suas concomitantes na outra língua a partir de um corpus previamente traduzido.

Charniak (1996) explica a aprendizagem de língua estatística de modo bastante acessível para linguistas, uma vez que ele advém de uma formação da matemática computacional. Para ele, essa metodologia aprende uma língua por intermédio de um corpus ao notar suas regularidades. Entretanto, o autor também afirma que é preciso adicionar um tanto de conhecimentos de língua, bem como de mundo para que os programas encontrem as regularidades corretas. Segundo Somers (2003), tanto a *Rule-Based System* quanto a abordagem estatística foram parcialmente bem-sucedidas até agora. Assim, modelos híbridos, que unem os dois tipos surgiram. Isso porque há problemas específicos de uma abordagem que a outra não é capaz de solucionar. Por exemplo, a resolução da anáfora é mais bem realizada pela *Rule-Based System* porque faz o “raciocínio” linguístico quando incorporada de uma ontologia, enquanto a estatística incorre nesse tipo de erro com mais frequência. Em contrapartida, o léxico específico de uma área do conhecimento teria mais sucesso na abordagem estatística, devido à frequência do número de ocorrências entre o grupo de palavras mais próximo (SOMERS, 2003).

5. Comparação dos tradutores do poema *A Pera*

O poema *A Pera* está na antologia poética de Vinícius de Moraes, composto em Los Angeles em 1947. Das várias camadas de leitura do poema a que está mais à vista é a descrição da imagem de uma fruta, a pera, retratada em natureza morta. No *Quadro 1: Traduções*, a seguir, há o poema original escrito pelo poeta e duas traduções, uma feita pelo tradutor do Google, considerado atualmente o mais eficiente em traduções automáticas e de fácil acesso, e outra tradução feita por Ashley Brown. A tradução feita por Brown foi encontrada no *website* do ex-professor Antônio Miranda.

Quadro 1: Traduções

Poema original em português	Poema traduzido pelo Google	Poema traduzido por Ashley Brown
A PERA Como de cera E por acaso Fria no vaso A entardecer A pera é um pomo Em holocausto À vida, como Um seio exausto Entre bananas	THE PEAR As wax And by accident Cold in the vase At dusk The pear is a doorknob Holocaust To life as An exhausted breast Among bananas	THE PEAR As if of wax And by chance Cold in the dish Growing late The pear is a fruit Burnt offering To life, like A breast exhausted Among bananas

Supervenientes E maçãs lhanas	Supervenient And orange apples	Extraneous And apples, candid
Rubras, contentes A pobre pera: Quem manda ser a?	Blush, happy The poor pear: Who is the boss?	Ruddy, content The poor pear: Who brings it to be?

Fonte: http://www.antoniomiranda.com.br/poesia_ingles/vinicius_de_moraes.html.

O poema é bastante metafórico ao comparar a pera com as outras frutas, ela é fria enquanto as outras são rubras; de acordo com Lakoff e Johnson (1980) o vermelho é calor. Outra analogia é a da pera como um seio exausto, porque de fato, a forma de uma pera, diferentemente de uma maçã, pode remeter à um seio envelhecido. Ainda, o retrato da pera em natureza morta é a própria metáfora da morte, expressa desde o início do poema em palavras como *cera*, tal um cadáver, e depois em holocausto, a chacina. Em termos de ambiguidade, o que está em jogo é a trama fonológica que inicia e encerra o poema entre a palavra *cera* e o amálgama *ser a*. O poeta não dá continuidade ao poema depois do artigo definido *a*, assim, parece ficar sob a responsabilidade do leitor preencher a lacuna *ser a quê?* A outra ambiguidade fonológica está no título em inglês *pear*, esta palavra é um homófono que remete tanto a fruta pera, bem como a palavra *par*, *casal*.

A ambiguidade em *cera* e *ser a* pode acontecer porque a frase *quem manda ser a?* pode ser entendida por, pelo menos, três sentidos. O primeiro deles é o mais aparente, o do verbo de estado *ser*, na expressão coloquial *quem manda ser assim?* como num tom de repreensão por se tratar de uma pera. Os outros dois sentidos de *quem manda ser a* pode ser compreendido pelo verbo *mandar* equivalente a *dar ordens*, assim, o significado pode ser *quem manda cera?* ou *quem manda, será?*. Nessas duas situações os fonemas *ce-ra* e *se-rá* mal se distinguem e são diferenciados pelos grafemas. Na opção *cera*, a pergunta é quem é o mandante da morte, já que a *cera* é uma metáfora do cadáver. Na segunda opção há uma indagação do poeta em *quem manda, será?* porque remete à uma dúvida no jogo de poder entre as frutas do vaso, qual delas manda? É possível que outras interpretações criem mais ambiguidades, pois, conforme dito anteriormente, há camadas de leitura no poema, aqui, entretanto, lidamos com as três hipóteses descritas acima.

Uma vez descrita a ambiguidade fonológica do poema *A Pera*, é possível analisar a tradução feita pela máquina e por humanos, que é o objeto deste estudo. Na tradução do *Google* o verbo *mandar* ganhou destaque, já que a relação semântica se estabelece entre mandar e *boss*, daí a pergunta quem é o chefe; uma inferência de quem manda ou dá as ordens. Na tradução feita por Brown, o raciocínio parece manter a transparência do verbo *ser* e indagar a identidade da pera, em *it to be*. A expressão como um todo *who brings it to be* também dá a ideia do questionamento sobre quem torna a pera ser quem é. A primeira pergunta quem é o mandante, já na segunda, pergunta-se quem deixa a pera assim? Há aproximações semântico-pragmáticas entre as duas traduções, mas elas não acarretam necessariamente o mesmo significado.

O que se pensa a partir dessas traduções é que a da máquina não utilizou um sistema de regras da língua, pois ao utilizar a palavra *boss/chefe*, sequer mencionada no poema, faz uma inferência (implicatura pragmática) pelo contexto do verbo mandar. Essa relação com o verbo foi possivelmente encontrada no corpus eletrônico,

em que traduções anteriores estavam armazenadas na base de dados, tudo o que a máquina faz é minerar esses dados. Já na tradução humana, ao manter o verbo *ser*, o tradutor parece tentar ser mais fiel a tradução original. Ainda que a tradução da máquina envolva somente dados estatísticos, ela parece interpretar o dito, ao passo que na tradução humana há uma preocupação com a fidelidade do texto. De fato, a tradução feita pela máquina também conta com traduções realizadas pelos humanos, pois são essas as informações que o minerador do Google encontra. Na seção a seguir, que encerra o artigo, faz-se um fechamento sobre as discussões da comparação entre os dois tipos de tradução.

6. Discussão

Este trabalho teve o intuito de analisar a desambiguação lexical realizada pelos tradutores humanos e as máquinas, área da linguística computacional chamada de *machine translation*. Há uma óbvia diferença entre os dois tipos de tradução, em primeiro lugar porque humanos têm consciência linguística que repercute também em consciência tradutória que a máquina não tem. Os tipos de raciocínio feitos durante a tarefa cognitiva de desambiguar o léxico durante os dois tipos de tradução foram os assuntos discutidos nas primeiras seções do artigo. Na última seção comparou-se as traduções feitas do poema *A Pera* entre o Google e A. Brown, e observou-se que nas duas há o quesito fidelidade a obra original, mas o Google realizou uma implicatura pragmática ao encerrar o poema que acabou por diferenciar as duas traduções.

A escolha pela poesia para um estudo sobre desambiguação lexical aconteceu porque este gênero textual é bastante metafórico e frequentemente apresenta desvios de linguagem que podem causar dificuldades para os tradutores. A ambiguidade estudada no poema foi, de fato, a partir da fonologia entre as palavras *cera* e *ser a*. É na pronúncia entre uma e outra (quase um homófono *cera/sera*) que o tradutor precisa desambiguar, contudo, esse jogo ambíguo das palavras não é óbvio, porque os grafemas são diferentes e a trama fonológica pode passar despercebida pelo leitor ou tradutor menos consciente.

A pergunta inicial de como os tradutores desambigam é respondida aqui pela comparação do escopo teórico tanto da área de tradução humana de um modo geral, (BEVILACUA; REUILLARD, 2013) quanto pela pesquisa em inteligência artificial, especificamente, *machine translation* (MITKOV, 2003; HUTCHINS, 2003). Enquanto o tradutor humano precisa ter domínio dos subsistemas das línguas de partida e alvo e do desenvolvimento de uma consciência tradutória que envolve, inclusive, estilística e criatividade linguística, a tradução pela máquina tanto conta com uma programação dos sistemas de regras de uma língua quanto com uma base de dados estatística na qual a frequência de uso de palavras já traduzidas indica qual a melhor tradução. Enquanto a tradução humana necessita das próprias capacidades cognitivas, tal como a consciência, a máquina lida com dados numéricos, já que funciona através de programação computacional. Essa diferença entre uma e outra parece ser uma obviedade, mas, de fato, o avanço dos tradutores eletrônicos, principalmente baseados em sistemas híbridos, mostra como esse tipo de metodologia tem funcionado.

Quanto aos resultados da comparação entre um tipo de tradução e outro o que se avalia é que a tradução feita pelo Google foi muito próxima da tradução humana.

Isso porque o poema já havia sido possivelmente traduzido e a mineração dos dados foi potente o suficiente para extraí-los. Não há um pensar linguístico *per se* nessa tradução, se há criatividade linguística ela está muito mais voltada para a criatividade da linguagem de programação computacional do que da criatividade linguística que utilizamos ao traduzir. Nesse sentido, a característica da originalidade quando a máquina traduz é o que impede melhores traduções diante de textos novos. No caso de um poema com desvios de linguagem, tal como metáforas novas ou ambiguidades de palavras não frequentes que não estão em quantidade suficiente na base de dados, então a máquina possivelmente terá dificuldades.

Referências

BEVILACQUA, C. R.; REUILLARD, P. C. A formação em Tradução na Universidade Federal do Rio Grande do Sul. In: GUERINI, A.; TORRES, M-H. C.; COSTA, W. (orgs.). *Os Estudos da Tradução no Brasil nos Séculos XX e XXI*, 2013, p.121-134.

CHARNIAK, C. *Statistical language learning*. Cambridge: MIT Press, 1996.

M.DE LA ROSA, K. ESKENAZI, M. Impact of word sense disambiguation on ordering dictionary definitions in vocabulary learning tutors. In: 24ª Conferência da Sociedade de pesquisa em Inteligência Artificial. 2011, Florida, Anais da *Association for the Advancement of Artificial Intelligence* (www.aaai.org) . Flórida,EUA, 2011,p.507-12.

DEHEANE, S. *Os neurônios da leitura: como a ciência explica nossa capacidade de ler*. Porto Alegre: Penso, 2012.

FILLMORE. C. Frames and the semantics of understanding. *Quaderni di Semantica*.vol.6,n.2, p.222-54, 1985

JUFFS, A.; HARRINGTON, M. Garden path sentences and error data in second language sentence processing. *Language learning* 46, 2, 1996.

LAKOFF, G.; JOHNSON, M. *Metaphor we live by*. Chicago:The University of Chicago,1980.

LAUFER, B. Assessing the effectiveness of monolingual, bilingual and 'biligualised' dictionaries. *The Modern Language Journal*, 81, 2, 1997.

LESK, M. Automatic sense disambiguation using machine readable dictionaries : how to tell a pine cone from an ice cream cone. *Proceedings of AGMSIGDOC Conference.Toronto*, 1986.

MCKEOWN, M.; BECK, I.; KUCAN, L. *Bringing words to life: robust vocabulary instruction*. New York,Guilford Press. 2002.

MORAES, V. A *pera*. disponível em: http://www.antoniomiranda.com.br/poesia_ingles/vinicius_de_moraes.html. Acesso em 29 Nov. 2018.

MITKOV, R. (ed.). *The Oxford handbook of computational linguistics*. New York, Oxford Press, 2003.

PERFETTI, C. *Reading Ability*. Nova Iorque: Oxford University Press, 1985.

_____. *Reading skills: Lexical Quality for comprehension*. Scientific studies of Reading, vol.11 n.4, p.557-83, 2007.

PINTO, G.; WANNMACHER, V. Apresentação: Tecnologias, procedimentos de pesquisa e aplicações à linguagem e à comunicação numa perspectiva psicolinguística. *Letras de Hoje*, Vol. 50, nº. 1, 2012.

RUMELHART, D. E. Schemata: the building blocks of cognition. *In: SPIRO, R.J. (eds) Theoretical Issues in Reading Comprehension*. Hillsdale, New Jersey: Lawrence Erlbaum, 1980.

SOMERS, H. Machine translation: latest developments. *In: MITKOV, R. The Oxford handbook of computational linguistics*. Oxford: Oxford University Press, 2003.

STANOVICH, K. E. Matthew effects in reading: some consequences of individual differences in the acquisition of literacy. *Reading research quarterly*, Vol. 22, p. 360-407, 1986.

STEVENSON, M.; WILKS, Y. Word sense disambiguation. *In: MITKOV, R. The Oxford handbook of computational linguistics*. Oxford: Oxford University Press, 2003.

WALTZ, D.; POLLACK, J. Massively parallel parsing: a strongly interactive model of natural language interpretation. *Cognitive Science*, vol.9, n.1, p.51-74, 1985.